Available online at www.sciencedirect.com
ScienceDirect
Physics of Life Reviews 4 (2007) 189–222
PHYSICS of LIFE reviews
www.elsevier.com/locate/plrev

Review

Seeing surfaces: The brain's vision of the world

Heiko Neumann [a], Arash Yazdanbakhsh [b,c], Ennio Mingolla [b,*]

[a] *Ulm University, Faculty of Engineering and Computer Sciences, Institute of Neural Information Processing, Germany*
[b] *Boston University, Department of Cognitive and Neural Systems, USA*
[c] *Harvard Medical School, Neurobiology Department, USA*

Received 9 August 2007; received in revised form 4 September 2007; accepted 4 September 2007

Available online 19 September 2007

Communicated by L. Perlovsky

Abstract

Surfaces of environmental objects are the key to understanding the visual experience of primates. Surfaces create structure in patterns of light available for sampling by visual systems, and delineate potential interactions that an animal can have with its environment, such as approaching goals, avoiding obstacles, grasping an object, or identifying members of a social group. Recent progress in modeling the perception of visual surfaces highlights the importance of feedforward and feedback connections in visual neural networks that segregate and group visual input into coherent regions related to corresponding surfaces in the visual world. Rich non-linear network dynamics in the brain underlie surface perception, including the detection, regularization, and grouping of visual boundaries between surfaces, the determination of "ownership" of a boundary by a closer surface that partially occludes a background, and the apprehension of a surface's visual quality, such as color or texture. Recent modeling efforts on these fronts are reviewed.

PACS: 87.19.L-; 84.35.+i; 87.19.lj; 87.19.lt

Keywords: Neural networks; Perception; Receptive field; Visual system

Contents

* Corresponding author.
E-mail addresses: heiko.neumann@uni-ulm.de (H. Neumann), Arash_Yazdanbakhsh@hms.harvard.edu (A. Yazdanbakhsh), ennio@cns.bu.edu (E. Mingolla).

1571-0645/$ – see front matter
doi:10.1016/j.plrev.2007.09.001

1. The surfaces of our visual world

A key visual competency of many species, including humans, is the ability to rapidly and accurately ascertain the sizes, locations, trajectories, and identities of objects in the environment. Whether noticing a deer moving behind a thicket, locating edible plants, or steering around obstacles to reach goals, many of the tasks of vision can be understood as stemming from the need to guide behavior based on changing visual input, whose patterns of light are structured by the geometries of surfaces and illuminants. The psychologist James J. Gibson [1–3] devoted enormous energy to articulate the importance of surfaces as the interface between light and objects. Gibson noted that the patterns of light are highly specific to the layout of surface in a particular environment and thus, in principle, informative for any animal's visual system. The difficulties begin, however, when one considers the virtually infinite variation in the patterns of visual input that follow from changes of an observer's viewpoint, or from motion of objects in the environment, or from disturbances in the patterns of illumination (e.g. from a passing cloud). The pattern of visual stimulation is constantly changing, whether from changes of perspective or progressive occlusion and disocclusion of objects, while our brains somehow (correctly!) perceive a stable underlying world through all the optical transformations.

The challenge of using patterns of light to perceive a world is evident to any one who has ever tried to do computer vision. More than a quarter of a century has passed since the publication of David Marr's Vision [4], a volume that was a manifesto for a generation optimistic that artificial intelligence and computational vision would soon unlock secrets that had eluded generations of philosophers and sensory psychologists. While tremendous progress on machine vision has certainly been made in the interim, the rate of progress has undershot expectations from the 1980s. Getting a machine to perform anything like what animals do—navigating and interacting with objects based on available light from unknown sources—remains an elusive challenge.

A key thesis of this review is that much of the bottleneck is due to an initial reluctance to integrate findings from perceptual psychology and, more recently, neurophysiology. As a result, many previous models have tried to go from simple elementary features, whether "raw" pixel values or simple blobs or edges, directly to such high-level competencies as object recognition. Typically, such models would work up to a point for relatively sparse or noise-free data sets, but fail to generalize to more complex environments. The increase in computer processing power has steadily increased the size of data sets that can be processed, but fundamental obstacles still exists, and many of these concern surface perception as the key intermediate step between patterns of light and the richness of visual experience.

This article will not address all the important aspects of even surface perception. Specifically, we will not address the important issue of surfaces in motion, beyond noting some rudimentary issues in the immediately following paragraphs. Instead we will focus on the case of static surface perception. We note in passing that there is no such thing as truly "static" perception of anything, as our eyes are constantly making small movements (tremor, drift, or micro-saccades) even if we are "fixating" on a stationary target in a room, but for analytical purposes it is useful to separate out such cases from others where objects or the observer are making macroscopic movements. As we will describe, significant progress has been made in recent years in both experiments and modeling of surface perception.